

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-331407

(P2001-331407A)

(43)公開日 平成13年11月30日 (2001.11.30)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テ-マコ-ト(参考)
G 06 F 13/00	5 5 0	G 06 F 13/00	5 5 0 L 5 D 0 1 5
3/16	3 4 0	3/16	3 4 0 N 5 D 0 4 5
G 10 L 13/00		G 10 L 3/00	Q
15/00			5 5 1 P
15/28			5 5 1 A

審査請求 未請求 請求項の数14 O.L (全 19 頁)

(21)出願番号	特願2000-306406(P2000-306406)	(71)出願人	591102095 三菱スペース・ソフトウェア株式会社 東京都港区浜松町2丁目4番1号
(22)出願日	平成12年10月5日 (2000.10.5)	(71)出願人	500466566 ジエイジー・リンク・インク J Z L i n k, I n k.
(31)優先権主張番号	09/574990		アメリカ合衆国、カリフォルニア州、ロ
(32)優先日	平成12年5月19日 (2000.5.19)		ス・ガトス、ウォーセスター・ループ
(33)優先権主張国	米国 (U.S.)		128
		(74)代理人	10005/874 弁理士 曾我 道照 (外6名)

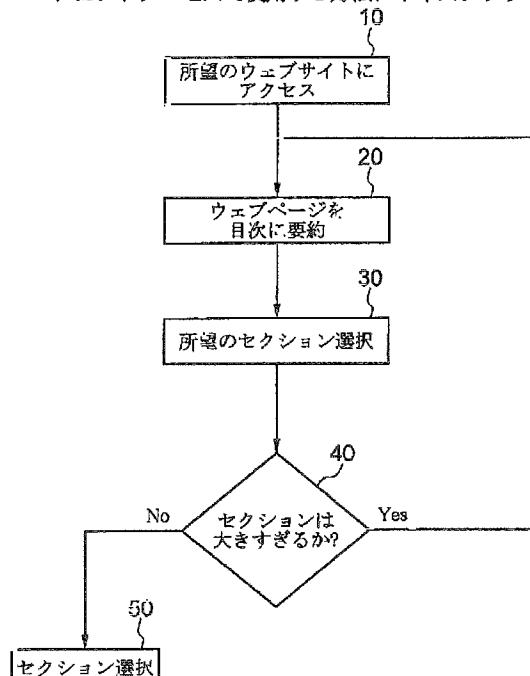
最終頁に続く

(54)【発明の名称】 ユーザがアクセス可能なウェブページを変換する方法、ユーザがアクセス可能なウェブページを自動的に変換するためのインテリジェントエージェントプロセスを使用する方法、ボイスプラウ

(57)【要約】

【課題】 ウェブサイトを音声フレンドリにすることが可能なボイスプラウザおよび変換システムを提供する。

【解決手段】 ボイスプラウザおよび変換システムは、ウェブページを読み上げるためのテキスト／スピーチ変換ソフトウェアと、単純なキーストロークまたはスピーチ認識によって次のページへナビゲートするための機構を含む。エンドユーザは、基本的な電話機を介してコンピュータへのアクセスを要求し、かつコンピュータネットワークをナビゲートでき、エンドユーザがページにマスクを作成することにより、カスタマイズされたニュースおよび他の情報を検索し特定の要求したウェブページにアクセスできる。各エンドユーザは、スポーツ、天気等の主要な情報を含みうる個人用のページをカスタマイズでき、電話機またはパーソナルデジタルアシスタント等の他のデバイスを介して独自の個人的なページにアクセスできる。



【特許請求の範囲】

【請求項1】 ユーザがアクセス可能なウェブページを変換する方法であって、  
所望のウェブページにアクセスするステップと、  
前記ウェブページを、複数のセクションを備える目次に要約するステップと、  
前記目次から所望のセクションを選択するステップと、  
該所望のセクションが大きすぎるフォーマットで提示されている場合、前記要約するステップを繰り返すステップと、  
前記所望のセクションが大きすぎるフォーマットで提示されていない場合、該所望のセクションを選択するステップと、  
前記選択されたセクションを表示するステップとを含む、ユーザがアクセス可能なウェブページを変換する方法。

【請求項2】 ユーザがアクセス可能なウェブページを自動的に変換するためのインテリジェントエージェントプロセスを使用する方法であって、  
ウェブページの統計を計算するステップと、  
前記計算された統計を所定の統計のセットと比較するステップと、  
前記計算された統計が前記所定の統計のセットよりも大きい場合、前記ウェブページをさらに分割するステップと、  
前記計算された統計が前記所定の統計のセットよりも大きくなる場合、前記ウェブページを再生用に翻訳するステップとを含む、ユーザがアクセス可能なウェブページを自動的に変換するためのインテリジェントエージェントプロセスを使用する方法。

【請求項3】 前記ウェブページの統計を計算するステップは、前記ウェブページの統計を要約すること、および最初に重要なリンクをリストした目次を作成すること、をさらに含む、請求項2記載の方法。

【請求項4】 前記重要なリンクは、前記ウェブページにおける該リンクの位置に基づいて決定される、請求項3記載の方法。

【請求項5】 クッキー(cookie)をバスするステップをさらに含む、請求項2記載の方法。

【請求項6】 コンピュータネットワークと通信するよう適合されたウェブサイトサーバと、  
前記コンピュータネットワークと通信するための通信手段を備えるネットワークゲートウェイサーバと、  
前記ネットワークゲートウェイサーバに連結され、該ネットワークゲートウェイサーバを介して前記コンピュータネットワークと通信するよう適合されたボイスプラウザクライアントコンピュータと、  
前記ボイスプラウザクライアントコンピュータに連結され、ユーザによる前記コンピュータネットワークへのアクセス要求を受信するよう適合された電話リンクであつ

て、前記ユーザは前記コンピュータネットワークをナビゲートして、基本的なコンピュータネットワーク情報、または情報の編集されていない可聴バージョンへのアクセスを取得する、電話リンクと、  
前記ボイスプラウザクライアントコンピュータに連結され、前記コンピュータネットワークに含まれる情報を再生用に翻訳するよう適合された、変換器とを備える、ボイスプラウザおよび変換システム。

【請求項7】 前記コンピュータネットワークは、インターネットをさらに備える、請求項6記載のボイスプラウザおよび変換システム。

【請求項8】 前記コンピュータネットワークは、インターネットをさらに備える、請求項6記載のボイスプラウザおよび変換システム。

【請求項9】 ウェブページにエンドユーザがカスタマイズしたマスクを作成する方法であって、  
所望のウェブページにアクセスするステップと、  
対応する新しい画像が現れるように、前記ウェブページの特定部分を選択するステップと、  
所望の情報の近くにあるボックスをクリックすることにより、マスクして変換したいと望む情報を選択するステップと、  
該所望の情報の選択を確認するステップとを含む、方法。

【請求項10】 所望の情報を前記ウェブページ上に現れる項目の前後に注釈文を挿入するステップをさらに含む、請求項9記載の方法。

【請求項11】 カスタマイズされたコンテンツ選択である、前記所望の情報を保存するステップをさらに含み、次に前記所望のウェブページにアクセスするとき、前記アクセスしたウェブページからの実際の情報と、前記カスタマイズされたコンテンツ選択の両方を受信する、請求項9記載の方法。

【請求項12】 前記所望の情報の選択を確認するステップは、適切な編集タスク機能バーをクリックすることをさらに含む、請求項9記載の方法。

【請求項13】 前記所望の情報の選択を確認するステップは、カスタマイズする特定のコンテンツの選択に、「キーワード」を使用することをさらに含む、請求項9記載の方法。

【請求項14】 コンピュータネットワークと通信するよう適合されたウェブサイトサーバと、  
前記コンピュータネットワークと通信するための通信手段を備えるネットワークゲートウェイサーバと、  
前記ネットワークゲートウェイサーバに連結され、該ネットワークゲートウェイサーバを介して前記コンピュータネットワークと通信するよう適合されたボイスプラウザクライアントコンピュータと、  
前記ボイスプラウザクライアントコンピュータに連結され、ユーザによる前記コンピュータネットワークへのア

クセス要求を受信するよう適合された通信回線リンクであって、前記ユーザは前記コンピュータネットワークをナビゲートして、所望のコンピュータネットワークウェブページへのアクセスを取得する、通信回線リンクと、プロセッサと、

前記プロセッサにウェブページを自動的に変換するインテリジェントエージェントプロセスを実行させるための命令を格納するプログラムメモリとを備え、

前記命令は、

ウェブページの統計を計算するステップと、該計算された統計を所定の統計のセットと比較するステップと、該計算された統計が前記所定の統計のセットよりも大きい場合、前記ウェブページをさらに分割するステップと、前記計算された統計が前記所定の統計のセットよりも大きくない場合、前記ウェブページを再生用に翻訳するステップとを含む、ボイスプラウザおよび変換システム。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】この発明は、インターネット等のコンピュータネットワークにゲートウェイを提供する通信システムに関し、特に、エンドユーザが基本的な電話機を使用して、コンピュータネットワークに一般的にアクセスできるボイスプラウザおよび変換システムに関する。

##### 【0002】

【従来の技術】ワールドワイドウェブコンピュータネットワーク（すなわち、「インターネット」）は、過去数年間にわたって並外れた成長を達成してきた。インターネットは、現在、天気予報、交通情報、および映画リストのような単純な日常情報から、科学技術の最前線での話題まで、広範な情報を提供している。

【0003】従来のウェブ機構は、情報にアクセスするための手段を提供するため、コンピュータのフロントエンドを使用したウェブプラウザを介して情報を提供する。標準的なウェブプラウザは、様々なインターネットウェブページに書かれている、既存の言語をベースとしている。使用される最も一般的なプログラミング言語は、ハイパーテキストマークアップ言語（すなわち「HTML」：Hypertext Markup Language）である。プログラムは、単純な命令をHTMLで書いて、特定の画像、テキスト、ファイル等の検索を呼び出す。あるウェブページから別のウェブページに移動するためには、異なるウェブページをリンクする「ハイパーリンク：hyperlink」が使用される。視覚主導型の標準的な閲覧機構では、リンクが異なる色、通常青色でマークされ、リンクへの接続は、リンクを「クリック」することにより確立する。ウェブコンテンツのオーディオ伝送用に特に開発されたプログラミング言語である音声用拡張マークア

ップ言語（すなわち、「Voice eXtensible Markup Language」：Voice eXtensible Markup Language）等の他のプログラミング言語が存在するが、HTMLは、インターネットプログラミングの大部分の言語を構成している。

##### 【0004】

【発明が解決しようとする課題】インターネットは、最も一般的に、ユーザにより、HTMLの可視強調性と合致するコンピュータディスプレイを介してアクセスされる。しかし、視覚的表示は、ますます増大している広範なインターネット情報へのアクセスに対する需要を満たすには十分ではない。これらの需要を満たすために、テレビおよびハンドヘルド式デバイスの使用などの、ウェブプラウザをよりアクセス可能なものとする異なるアプローチがある。しかし、これらのアプローチは、キーパッドおよび／またはポインティングデバイスを備えた表示装置を必要とする、ウェブプラウザの基本的機構を変えるものではない。当技術分野では、テキスト情報の提示を制御するためのスタイルシートも周知であるが、最も存在しているスタイルシートは、テキスト情報の視覚的提示に制限され、ページを見ることができないユーザの助けとはならない。スタイルシートアプリケーションの例は、ワードプロセッサおよびデスクトップパブリッシング（DTP）において見出すことができる。テキストおよびスクリーンリーダもまた当技術分野で周知であり、このようなリーダは、音声合成を通してテキスト情報をオーディオで提示するために使用されている。しかしながら、HTML言語は元々視覚的表示のために書かれたものであるため、HTMLの単純な「翻訳」を発話すると、その「翻訳」は理解しにくく、かつ潜在的に、実質的な情報を誤って伝達する可能性がある音声エンドプロダクトに終わってしまう。

【0005】インターネットは、主に視覚ベースのアクセスおよびナビゲーション用に設計されているが、聴覚的に、または視覚をベースとして聴覚を主とした情報へのアクセスを要求するか、あるいはそれを好むインターネットユーザのグループが増大している。このようなユーザとしては、視覚障害者や、或いは、タイミング、ロケーション、または装置の近さを好みどおりにして、普通の電話機を介してインターネットにアクセスしたいと思む個人が含まれる。このため、インターネットのコンテンツに聴覚的にまたは音声機器のみを介してアクセスする個人に対するエンドユーザ経験を豊かにすることが望ましい。

【0006】したがって、インターネットへの一般的なアクセスを可能にするとともに、パーソナルデジタルアシスタント（「PDA」：Personal Digital Assistant）または専用の機能を持った携帯電話を必要とせずに、使いやすいボイスプラウザおよび変換システムが必要とされる。

【0007】この発明は、エンドユーザが標準的な固定

電話または携帯電話を介してユーザがインターネットに一般的にアクセスできるシステムを提供する。

【0008】

【課題を解決するための手段】この発明に係るユーザがアクセス可能なウェブページを変換する方法は、所望のウェブページにアクセスするステップと、前記ウェブページを、複数のセクションを備える目次に要約するステップと、前記目次から所望のセクションを選択するステップと、該所望のセクションが大きすぎるフォーマットで提示されている場合、前記要約するステップを繰り返すステップと、前記所望のセクションが大きすぎるフォーマットで提示されていない場合、該所望のセクションを選択するステップと、前記選択されたセクションを表示するステップとを含むものである。

【0009】また、この発明に係るユーザがアクセス可能なウェブページを自動的に変換するためのインテリジェントエージェントプロセスを使用する方法は、ウェブページの統計を計算するステップと、前記計算された統計を所定の統計のセットと比較するステップと、前記計算された統計が前記所定の統計のセットよりも大きい場合、前記ウェブページをさらに分割するステップと、前記計算された統計が前記所定の統計のセットよりも大きくなる場合、前記ウェブページを再生用に翻訳するステップとを含むものである。

【0010】また、前記ウェブページの統計を計算するステップは、前記ウェブページの統計を要約すること、および最初に重要なリンクをリストした目次を作成することをさらに含むものである。

【0011】また、前記重要なリンクは、前記ウェブページにおける該リンクの位置に基づいて決定されるものである。

【0012】また、クッキー(cookie)をバスするステップをさらに含むものである。

【0013】また、この発明に係るボイスプラウザおよび変換システムは、コンピュータネットワークと通信するよう適合されたウェブサイトサーバと、前記コンピュータネットワークと通信するための通信手段を備えるネットワークゲートウェイサーバと、前記ネットワークゲートウェイサーバに連結され、該ネットワークゲートウェイサーバを介して前記コンピュータネットワークと通信するよう適合されたボイスプラウザクライアントコンピュータと、前記ボイスプラウザクライアントコンピュータに連結され、ユーザによる前記コンピュータネットワークへのアクセス要求を受信するよう適合された電話リンクであって、前記ユーザは前記コンピュータネットワークをナビゲートして、基本的なコンピュータネットワーク情報、または情報の編集されていない可聴バージョンへのアクセスを取得する、電話リンクと、前記ボイスプラウザクライアントコンピュータに連結され、前記コンピュータネットワークに含まれる情報を再生用に翻

訳するよう適合された、変換器とを備えるものである。

【0014】また、前記コンピュータネットワークは、インターネットをさらに備えるものである。

【0015】また、前記コンピュータネットワークは、インターネットをさらに備えるものである。

【0016】また、この発明に係るウェブページにエンデューザがカスタマイズしたマスクを作成する方法は、所望のウェブページにアクセスするステップと、対応する新しい画像が現れるように、前記ウェブページの特定部分を選択するステップと、所望の情報の近くにあるボックスをクリックすることにより、マスクして変換したいと望む情報を選択するステップと、該所望の情報の選択を確認するステップとを含むものである。

【0017】また、所望の情報を前記ウェブページ上に現れる項目の前後に注釈文を挿入するステップをさらに含むものである。

【0018】また、カスタマイズされたコンテンツ選択である、前記所望の情報を保存するステップをさらに含み、次に前記所望のウェブページにアクセスするとき、前記アクセスしたウェブページからの実際の情報と、前記カスタマイズされたコンテンツ選択の両方を受信するものである。

【0019】また、前記所望の情報の選択を確認するステップは、適切な編集タスク機能バーをクリックすることをさらに含むものである。

【0020】また、前記所望の情報の選択を確認するステップは、カスタマイズする特定のコンテンツの選択に、「キーワード」を使用することをさらに含むものである。

【0021】また、この発明に係るボイスプラウザおよび変換システムは、コンピュータネットワークと通信するよう適合されたウェブサイトサーバと、前記コンピュータネットワークと通信するための通信手段を備えるネットワークゲートウェイサーバと、前記ネットワークゲートウェイサーバに連結され、該ネットワークゲートウェイサーバを介して前記コンピュータネットワークと通信するよう適合されたボイスプラウザクライアントコンピュータと、前記ボイスプラウザクライアントコンピュータに連結され、ユーザによる前記コンピュータネットワークへのアクセス要求を受信するよう適合された通信回線リンクであって、前記ユーザは前記コンピュータネットワークをナビゲートして、所望のコンピュータネットワークウェブページへのアクセスを取得する、通信回線リンクと、プロセッサと、前記プロセッサにウェブページを自動的に変換するインテリジェントエージェントプロセスを実行させるための命令を格納するプログラムメモリとを備え、前記命令は、ウェブページの統計を計算するステップと、該計算された統計を所定の統計のセットと比較するステップと、該計算された統計が前記所定の統計のセットよりも大きい場合、前記ウェブページ

をさらに分割するステップと、前記計算された統計が前記所定の統計のセットよりも大きくなる場合、前記ウェブページを再生用に翻訳するステップとを含むものである。

#### 【0022】

【発明の実施の形態】この発明は、エンドユーザが標準的な固定電話または携帯電話を介してユーザがインターネットに一般的にアクセスできるシステムを提供するもので、電話システムを使用して、ウェブページを聞くことにより、インターネットにアクセスする。インターネットをナビゲートするためには、女性の声等の明瞭な聲で発せられる言葉を繰り返すか、または電話機のボタンを押下するか、あるいは音声によるコマンドを与えて、リンクを選択し、情報をアクセスする。したがって、公衆電話機および/または携帯電話等の電話システムは広く使用されているため、インターネット情報は、エンドユーザにとってよりアクセス可能である。

【0023】また、この発明のボイスプラウザおよび変換システムは、ウェブページを読み上げるためのテキスト/スピーチ変換ソフトウェア(text-to-speech software)と、単純なキーストロークまたはスピーチ認識によって次のページへナビゲートするための機構とを含む。本システムは、概して、HTMLコードをベースとし、これを使用するが、特にスピーチまたはオーディオによる提示のために開発されたVoiceXMLとも互換性がありうる。

【0024】また、この発明のシステムは、エンドユーザに対して、(i) エンドユーザが個人用にカスタマイズし制御できる、特定の基本的なニュースヘッドライン、金融情報、天気情報等、および(ii) ウェブページの実質的に編集されていない「オーディオのみ」のバージョンを提供する。すなわち、本システムにより、エンドユーザは、カスタマイズしたニュースおよび他の情報の検索と、最初はエンドユーザによってカスタマイズされていないか、または編集されていない、エンドユーザが要求する特定のウェブページへのアクセスとの両方を行うことが可能である。このために、この発明のシステムは、(可能な最大程度まで)編集されていないHTMLベースのウェブページのオーディオのみのバージョンをエンドユーザに提供する。

【0025】特定のウェブページについては、本システムは、ページを変換/マスクして、ユーザが電話でアクセスできる独自のページを提供する。該サービスは、各顧客に対して個人用のページと、従来の視覚フレンドリ(visual friendly)なHTMLウェブページを聴覚フレンドリなウェブページに翻訳するサービスとを提供する。各エンドユーザは、電子メール、カレンダ、ニュース、スポーツ、天気、交通情報、映画リスト、およびその他のコンテンツ等の主要な情報を含みうる個人用のページをカスタマイズすることができる。

【0026】個人用のページは、「変換」または「マスキング」と呼ばれるプロセスによって、聴覚的または「音声フレンドリ」な状態にフォーマット化される。マスクは、エンドユーザが定期的にアクセスする各ウェブサイト毎に作成することが可能なため、エンドユーザが各自のリストニング優先順位でウェブページをカスタマイズすることが可能であり、これにより、ウェブページをレンダリングする様式を制御する。この目的のために、音声スタイルシートは、音量、速度、ピッチ、指示、特定要素についての出力抑制、文字によるテキスト文字のスペルアウト、スピーチフォント(男性/女性、成人/子供、等)、要素コンテンツ前後のテキスト挿入、音楽または予め録音したサウンドを含む、要素前後、および要素中、要素後のサウンド効果、等の特徴の範囲を制御するためのベースを提供することができる。

【0027】HTMLベースのボイスプラウザおよび変換システムのより完全な理解、ならびにそのさらなる利点および目的の認識は、以下の好ましい実施の形態の詳細な説明を考慮することで、当業者によりなされよう。添付図面が参照される。

【0028】まず、図1を参照して、この発明の一実施の形態による、ユーザが電話機を介してインターネットにアクセス可能なボイスプラウザシステムが提供される。ボイスプラウザシステムは、従来のインターネット技術を音声およびスピーチ処理用途にまで拡張するとともに、インタラクティブな音声応答技術の新しい使用を可能にする。ボイスプラウザシステムは、ウェブサイトサーバ3と、ボイスプラウザクライアントコンピュータ(VBCC)2と、インターネットゲートウェイサーバ(IGS)4と、インターネットゲートウェイ5と、電話リンク6とを含む。なお、ボイスプラウザシステムは、多数のウェブサイトサーバ、VBCC、IGS、および電話リンクを有してもよいが、図を整頓するために、それぞれ1つのみを図1に示すことに留意されたい。VBCC2は、アクセス要求に応答して、IGS4およびインターネット1を介してウェブサイトサーバ3にアクセスするための通信手段を備える。ウェブサイトサーバ3およびIGS4は、ハイパーテキスト転送プロトコル(HTTP)にしたがってインターネット1を介してファイルを転送する。VBCC2は、イーサネット(登録商標)を介してIGS4に接続されてもよく、IGS4はまたインターネット1を介してインターネット1に接続される。さらに、インターネット6は、アクセス要求を生成するため、電話線5とVBCC2の間の通信用に使用される。

【0029】電話機キーパッドエントリ(または回転式ダイヤル電話ではダイアル)および/または音声コマンドは、VBCC2へのエントリのために変換される。なお、アクセス要求を生成するためのGPRS(ラジオ)等の他の手段が存在することに留意されたい。次に、オーディオファイル(すなわちウェブページ)が変換され

て、ユーザの電話機を介して再生される。このように、電話で、IGS4に接続されたVBCC2にダイアルする。IGS4は、ラインT1を介してインターネット1に接続されるとともに、プロキシサーバを実行して、実際のインターネットプロトコル(IP)アドレスなしで、VBCC2はインターネット1にアクセスさせられる。図1のT1ライン、DSLライン等を含む各種形態の接続線を採用してもよいことを理解されたい。インターネットにおいて、www.company.com等のウェブサイトが提供され、VBCC2上で実行されているボイスプラウザは、www.company.com等の所定のウェブサイトへのデフォルトURLを有する。

【0030】インターネットにアクセスするために「ダイアルアップ」する必要がある状況では、接続を確立するために、エンドユーザーはサービスプロバイダの市内電話番号またはフリーダイヤル番号をダイアルすることができる。他の状況では、接続は、「ダイアルアップ」が必要ないような専用線を介して行われる。主に初心者ユーザのために、第1ページがエンドユーザーに対して読み上げられる。第1ページは、音声を使用して閲覧するための一般的な使用説明を含む。そして、エンドユーザーは、各自のマスタホームページへの直接リンクを選択することができる。簡略化のために、初期登録プロセスは、標準的なウェブブラウザ機構、すなわちコンピュータを使用した処理についてセットアップされる。サービスは、各エンドユーザーが初期登録時に必要な情報を入力するよう要求する。

a. エンドユーザーがサービスにダイアルインするために使用する電話番号。これはIDとして使用される。ユーザが「契約」した電話機ではない電話機から通話する場合、識別の代替モードを使用してもよい。たとえば、ユーザは要求される電話番号をキー入力してもよく、または音声認識を用いて識別してもよい。

b. エンドユーザーが聞きたいと望む最初のウェブページのハイパーリンク選択。エンドユーザーが音声用のウェブページを持っていない場合、サービスはデフォルトページを提供できる。

c. エンドユーザーが登録を完了した後、サービスが利用可能となる。そして、サービスは、電話番号からエンドユーザーを識別することができ、かつユーザの個人用のウェブページを見つけることができる。エンドユーザーの個人用のページから、エンドユーザーは、後述するプロセスによってウェブをナビゲートできる。

【0031】この発明のボイスウェブブラウザおよび変換システムに関する3つの基本的な技術がある。これは、すなわち、インターネット技術と、電話通信技術と、音声合成および認識技術である。インターネット技術に関して、ウェブブラウザはウェブページのテキストおよび画像を視覚的に表示する代わりに、テキストを読み上げるよう変更される。httpサーバの変更が必

要ないように、標準的なhttpプロトコルを使用することもできる。この発明の一実施の形態において、システムは、エンドユーザー契約者に、実質的に編集されていない音声によるウェブページ、または表示形態を有する音声によるウェブページを提供する。エンドユーザーは、個々に選択したマスク(図8～図13に関連して詳細に後述する)を作成することで、音声によるウェブページまたは表示を有する音声によるウェブページをカスタマイズする能力を有する。

【0032】電話通信技術に関して、エンドユーザーの識別およびキーパッド入力等の電話入力は、サービスを提供するコンピュータと相互作用する。図1に関して示すインターフェース6の代替として、該技術の2つの主要部分が存在する。一方は電話サービスにリンクする交換システムである「構内交換設備(PBX)」であり、他方は、コンピュータと、電話要求を処理する「コンピュータ電話通信相互作用(CTI)」と呼ばれる1つのソフトウェアである。たとえば、Microsoftの電話通信アプリケーションプログラムインターフェース(TAPI)を使用してもよい。そして、ボイスブラウザは、エンドユーザーがボタンを押下して、ハイパーリンクの選択、後方へのページ移動、リンクリストの傾聴等を行うことを可能にする。また、エンドユーザーを識別するために、発信者IDをエンドユーザーから抽出してもよい。

【0033】音声合成および音声認識に関して、既存の合成アプリケーションプログラミングインターフェースは、ウェブブラウザコードに組み込むことができる。音声認識技術は、数個の「キーワード」から選択するためには使用される。現行技術は、限られた語彙の認識とは無関係に、発話者を適切にサポートしうる。ボタンの押下は最低限に抑えられる。

【0034】この発明の好ましい実施の形態の動作環境のような純粋なオーディオ状況では、エンドユーザーが、インターネットにアクセスする際に操作する視覚的なオブジェクトがない。また、情報フローは二次元ではなく、音声合成器が読み上げる一連のテキストである。異なるオーディオコンテンツを示すためには、異なるトーンまたは異なる背音を使用する。たとえば、リンクを示すには、女性の声のトーンがあてられ、通常のテキストを読み上げるには男性の声が使用される。リンクを確立するには、電話機上の指定されたボタンを、リンクが読み上げられている時間制限内に押下する。こうして、インターネットをナビゲートするには、女性の声で発話された言葉を繰り返すか、または電話機のボタンを押下して、リンクを選択する。

【0035】なお、好ましい実施の形態では、本システムは、音声認識プラットフォーム(エンドユーザーが電話で話すことにより、情報を入力、要求する)、電話キーパッドプラットフォーム(エンドユーザーが電話機のキーパッドを使用することにより、情報を入力、要求す

る）、または複合プラットフォーム（エンドユーザが、話すようプロンプトされたときには話すことにより、また特定のエントリには電話機のキーパッドを使用することにより、情報を入力、要求する）をサポートすることに留意されたい。

【0036】インターネットをナビゲートする例示的な実施の形態において、エンドユーザには、以下の一連の選択肢が与えられる。

詳細な使用説明については、1を押下。

ページ選択には、2を押下。

株式相場については、3を押下。

企業ニュースについては、4を押下。

法人ニュースについては、5を押下。

私の企業リンクについては、6を押下。

レストラン企業については、7を押下。

星占いについては、8を押下。

企業ログインについては9を押下。

【0037】例示的なナビゲーションプロセスにおいて、詳細な使用説明へのリンクを選択するには、下線の付いた語「使用説明」が、女性の声等の明瞭な音声で話され、エンドユーザがこれを繰り返して、そのリンクを選択する。あるいは、エンドユーザは、詳細な使用説明へのリンクを選択するために、「1」とマークされた電話機のボタンを押下してもよい。「企業ニュース」たとえば、Yahoo!、MSNBC、またはCNNの各ニュースサイトへのリンクを選択するには、下線の付いた「企業ニュース」という語が女性の声で話され、エンドユーザはその語を繰り返して、そのリンクを選択する。あるいは、エンドユーザは、企業ニュースへのリンクを選択するために、「4」とマークされた電話機のボタンを押下してもよく、以下同様である。

【0038】可能なリンクの例示的なリストの他に、エンドユーザは「カレンダ」を編集することができる。これを達成するために、上述したのと同じプロセスが続けられる。女性の声等の明瞭な音声が「カレンダ」という語を発し、エンドユーザがその語を繰り返すことによってリンクを選択することができる。あるいは、エンドユーザは、電話機上の適切なボタンを押下することで、リンクを選択してもよい。パーソナルカレンダを使用して、電話による通知を可能にできるため、エンドユーザの指定されたスケジュールを変更することができる。

【0039】リンクを確立するために、ボイスプラウザは、エンドユーザに電話機のキーパッド上のリンクに対応する番号ボタンを押下するよう求める。正方向および逆方向へのページアクセスは、指定されたボタン「0」および「#」でナビゲートされる。あるいは、ユーザは、ページを移動するために、「正方向」および「逆方向」というコマンドを使用してもよい。携帯電話用途において、キーパッド上のボタンを押下するために耳からハンドセットを離さなければならないのは面倒である。

したがって、音声認識は、たとえばエンドユーザにリンクの語を読み返すよう求めることにより、インラクティブに使用されてもよい。たとえば、エンドユーザは、次へと望むハイパーテキストリンクを聞いたときに、「次へ」という語を発することができる。エンドユーザはまた、ブラウザを中断して関連トピック（リンク）のショートリストを要求することも可能である。

【0040】基本的な音声認識技術は、エンドユーザがブックマークリストから「ブックマークされた」気に入ったサイトにアクセスし、かつ「ブックマーク追加」して、現在のページをリストに載せられるようにするために使用される。ブックマークを用いると、ユーザは、ユーザの個人用のページにリストされるリンクを含む個人的なHTMLファイルを保存することができる。

【0041】本システムは、エンドユーザに対して、(i) 特定の基本的なニュースヘッドライン、金融情報、天気情報等、および (ii) ウェブページの実質的に編集されていない「オーディオのみ」のバージョンを提供する。すなわち、ボイスプラウザシステムにより、エンドユーザは、カスタマイズしたニュースおよび他の情報の検索と、実際のウェブでの「サーフィン」、すなわち特定のエンドユーザが要求するウェブページへのアクセスと、の両方を行うことが可能である。このために、本システムは、(可能な最大程度まで) 編集されていないオーディオのみのバージョンのHTMLベースのウェブページをエンドユーザに提供する。

【0042】特定のウェブページについては、本システムは、ページを変換／マスクして、エンドユーザが電話でアクセスできる独自のページを提供する。本システムは、各顧客に対して個人用のページと、従来の視覚フレンドリなHTMLウェブページを聴覚フレンドリなウェブページに翻訳するサービスとを提供する。なお、大部分のウェブページの元のコンテンツはHTMLであるが、本システムは、HTMLからWML（画面を備える携帯電話で通常使用される）、Web-c1ippin g、i-m o d e（登録商標）等の他のフォーマットに、またはこれら他のフォーマットからHTMLに変換する能力を有する。各エンドユーザは、電子メール、カレンダ、ニュース、スポーツ、天気、交通情報、映画リスト、およびその他のコンテンツ等の主要な情報を含みうる個人用のページをカスタマイズすることができる。個人用のページは、「変換：transcoding」または「マスキング：masking」と呼ばれるプロセスによって、音声によるページまたは視覚的な表示を有する音声によるページにフォーマット化される。このプロセスによって、特に、電話機、PDA、または他の同様のデバイスによるアクセス時に、ウェブページをより簡潔にかつユーザフレンドリに表示することが可能になる。従来のスタイルシート言語は、注釈を追加できるほど柔軟ではないため、本システムは、マスクと呼ばれるHTML用の

独自の言語を使用する。マスクは、重要なウェブサイトそれぞれについて作成される。ページは、同様に視覚的に読むことができるよう保持されるべきである。変換は、スマートフォン、画面付き公衆電話等のオーディオおよびビジュアルシステムの双方と互換性があり、オーディオおよびビデオの各リンクを提供し、ウェブページを音声と同期させることができる。

【0043】エンドユーザーは、ドキュメントをどのようにレンダリングするかを制御できる。この目的のために、聴覚スタイルシートは、音量、速度、ピッチ、指示、特定要素についての出力抑制、文字によるテキスト文字のスペルアウト、スピーチフォント（男性／女性、成人／子供等）、要素コンテンツ前後のテキスト挿入、音楽または予め録音したサウンド（たとえば、海岸でくだける波の音）を含む、要素前後、および要素中のサウンド効果、等の特徴の範囲を制御するためのベースを提供することができる。

【0044】変換システムを使用する際に、3つの一般的な代替例がある。すなわち、(i) マスクが望ましくない場合には、後述するように、インテリジェントエージェント（Intelligent Agent）を使用する自動変換が一般に使用される。(ii) ユーザがカスタマイズしたマスク、すなわち、ユーザがYAHOO!、MSNBC、またはCNN等のウェブサイトに赴いて、視覚的編集ツールを使用して、ユーザに独自の、所望の部分のマスクを選択し、保存してもよい。また、(iii) コンテンツプロバイダ、たとえば、YAHOO!、MSNBC、またはCNNは、各自のマスクを作成してもよい。

【0045】ボイスブラウザおよび変換システムのさらなる利点としては、クッキー（cookie）に変換システムをバスさせる能力が含まれる。クッキーは、HTMLサイトならびにウェブサイトから適用されるため、電子商取引等のインタラクティブアプリケーションが本システムを通して容易になる。さらに、本システムは、企業のためのインターネットビジネスに適用されて、各自のインターネット特徴を強化することができる。

【0046】図2を参照して、この発明の好ましい実施の形態の自動変換が、ウェブページに対してどのように動作するかの流れ図を提供する。ステップ10において、適宜ウェブサイトアドレスを選択することで、ウェブサイトが選択される。説明の目的のため、COMPANYの元のページ、すなわちホームページを図3に示す。COMPANYウェブページ（すべてHTMLベースのウェブページ）は、一連のネストされた画面から構成される。すなわち、図3（通常ウェブページに含まれるグラフィックまたはピクチャを図示せず）を参照すると、「BREAKING NEWS」（概して12で示す）、「COMPANY SHOPPING」（概して14で示す）または「EDUCATION」（概して16で示す）等のリンクは、ユーザがこれらのリンクを選択するか、またはクリックすると、他の

ネストされた画面へのアクセスを提供する。YAHOO!、MSNBCまたはCNN等いずれの企業であってもよいCOMPANYのメインページは、<http://www.COMPANY.com>を選択することにより、選択、アクセスされ、リンクを選択することで、識別された様々な主題にアクセスする。図3に例示するように、「BREAKING NEWS」（12）のサイズを「Privacy Policy」（概して18で示す）と比較することでわかるように、リンクは様々なタイプのサイズを有してもよい。サイズは、相対的な重要性、話題等を含む任意の数の要因に基づきうる。

【0047】ステップ20において、エンドユーザーには、異なるリンクを示すショートタイトルにより提示される異なるセクションを含む目次の形態で要約されたウェブページが提供される。COMPANYウェブページの場合、目次は、「BREAKING NEWS」、「BUSINESS」、「Company Info」等のショートタイトルで示されるセクションを含みうる。ステップ30において、所望のセクションが選択される。たとえば、「BREAKING NEWS」セクション（図3では概して12で示される）が、その特定のリンクをクリックすることにより選択される。ステップ40において、そのセクションが大きすぎるか否かが決定される。図6に関して詳細に後述するように、セクションが大きすぎる場合、本システムは、ステップ20に戻り、要約プロセスを繰り返す。セクションが大きすぎない場合、本システムはステップ50に進み、適宜選択が行われる。たとえば、図3のウェブページを参照して、「BREAKING NEWS」セクションが、セクションのその部分をクリックすることで選択されると、ユーザはニュースに集中した別のウェブページに接続される（図4に示すように）。図4のウェブページは、通常様々な追加テキスト、グラフィックス、ピクチャ等、ならびにCOMPANYのその他のウェブページへのリンクも含むであろうことを理解されたい。セクション「COMPANY SHOPPING」を望む場合、セクションのその部分をクリックすることによりこのセクションが選択され、その結果、図5に示すようになる。ここでも、該ウェブページは、通常、ピクチャ、グラフィックス、リンク等を含んでもよく、また、図5において、識別するリンク等、リンクを選択することで様々な他のページにアクセスしてもよいことを理解されたい。

【0048】次に、図6を参照して、インテリジェントエージェントプロセスを例示する流れ図を提供する。この発明の一実施の形態において、HTMLベースのウェブページのオーディオ、またはビジュアル形態を有するオーディオへの自動変換は、「インテリジェントエージェント」プロセスを介して完了される。インテリジェントエージェントプロセスは、HTMLページを、本システムにのみ存在し、かつ元のウェブページから分割されたページである、ソフトページに自動的に翻訳するため使用される。ステップ210において、インテリジェ

ントエージェントプロセスは、COMPANYウェブページ等のHTMLウェブページのある「統計」を計算して、要約する。ステップ212において、プロセスはこれらの計算を所定の統計または基準のセットと比較する。この比較に基づいて、さらにページを分割するか否か、すなわち、次のネストされた画面に進むか、あるいはそこで停止して、ウェブページのネストされた画面のうちの1つを表示するか否かについての決定を行う。換言すれば、インテリジェントエージェントは、本システムがエリア内のリンク数、そのエリアが表を含むか否か、ノードの位置等の、ソフトページの統計を使用して、何を表示するかを決定する、比較プロセスを提供する。インテリジェントエージェントは、ネスト内に存在する情報が多すぎて、オーディオまたは視覚的な表示を有するオーディオのためにHTMLを変換または転換した(transcoding or converting)結果、見苦しい表示になるか否かを決定するよう試みるために、比較を行う。ステップ216において、ページの統計または基準が所定の基準を越える場合、表示が不適切であるものと仮定して、インテリジェントエージェントプロセスはさらにページを分割する。そして、プロセスは、ステップ210に戻り、ソフトページをさらに要約する。ステップ214において、ページの統計または基準が所定の基準よりも低い場合、インテリジェントエージェントプロセスは、画面ページが表示に適切であると仮定し、したがって、HTMLが変換される。表示自体に関して、変換は純粋なオーディオ表示またはオーディオ/ビジュアルが混ざった表示、たとえばその他の携帯電話またはPDAにおいて利用可能なもの等、の両方について達成することができる。

【0049】さらなる説明として、インテリジェントエージェントプロセスは、ウェブページの統計を要約して、サブセクションの統計、ページにおけるリンクの位置、およびページのその他任意の特徴において判定される、最も重要なリンクを最初に有する目次を作成する。たとえば、図3に示すCOMPANYのメインウェブページでは、インテリジェントエージェントは、最も重要なリンクは、「BREAKINGNEWS」(12)、「COMPANY SHOPPING」(14)等のように中間上部またはページの中間にあるものと仮定し、「Privacy Policy」(18)のようにページの底部にあるリンクは重要であるとはみなさない。比較プロセスの判断に基づいて、変換が完了し、HTMLテキストが再生のために変換される。なお、COMPANYページの底部にある「Privacy Policy」等の重要なリンクには依然としてアクセス可能でありうることに留意されたい。たとえば、ユーザは、デフォルトフレーズまたはキーを使用して、それが存在するか否かを見出すことができる。

【0050】次に、図7を参照して、COMPANYメインウェブページは、グラフィック画像なしであるが、

ページの上部にマスクエディタ80を有した状態で図示されている。マスクエディタ80は、エンドユーザがウェブページをカスタマイズするためにマスクを作成するための手段を提供する。マスクエディタ80はまた、図8～図13に関してより詳細に説明するように、エンドユーザが作成したテキストを挿入できる、ヘッダセクション82およびテールセクション84を含む。エンドユーザがマスクをカスタマイズできるようにするために、編集タスク機能86、88、90、92、94、および96が設けられる。図7に図示するように、マスクエディタが現れているときは、COMPANYページのグラフィック画像が省略され、ページが一連の縮小サイズ、すなわち「サブセット構成」ボックステキスト、たとえば98、100、102と、選択ボックス、たとえば104、106、および108にグループ化される。この構成は、HTMLが、各情報が一意に識別可能な位置またはノードに現れる本質的に「降順ツリー」言語であるため、提供されるものである。編集タスク機能、ボックステキスト、および選択ボックスの利用については、図8～図13に関して後述する。

【0051】図8～図13に目を向けて、エンドユーザによりカスタマイズされたマスクの実際の作成を検討する。図8は、企業からの典型的なウェブページ、たとえばwww.NEWSHOP.comの画像を提示する。これは、ウェブページにアクセスするために、コンピュータモニタシステムを使用して、エンドユーザに視覚的に提示される。幸いなことに、ある特定のHTMLコンテンツは、それが主にテキストで構成されているので、簡単に変換され、聴覚的に報知(音声出力)される。たとえばボックス110に含まれる言語がそうであり、グラフやチャートを含むボックス112の「Market Snapshot」のような他の部分はそうでない。図3～図5のウェブページのように、図8のウェブページは、通常ウェブページにおいて見られる典型的なピクチャ、リンク、グラフィックス、テキスト等を含まない。エンドユーザが、カスタマイズマスクを作成するページとしてこのページを選択すると、ウェブページの特定部分、Market Snapshot 112が選択されて、図9に図示するように新しい画像が現れる。

【0052】図9において、マスクエディタ80は、マスクが準備される実際の画像の識別/位置、またはHTMLファイルの降順ツリー構造に基づいた「ノードID」を含む。Market Snapshotチャートは、特定のテキストおよびテキストの直ぐ隣りに現れる、様々な空のボックス116、118、120、122、124、126を含むチャート114で表現される。図10において、エンドユーザは、所望のテキストの横にあるボックスをただクリックするか、あるいは、選択することにより、マスクされ変換された情報を選択する。図10において、これは、ボックス120、122、124、およ

び126における「チェックマーク」として例示される。選択されたボックスの精確なノードIDは、マスクエディタ80内に現れる。さらに、エンドユーザは、ヘッダテキストボックス128に情報をタイプすることで、ヘッダ82に所望のテキストを挿入する。図10に示すように、エンドユーザは、ヘッダテキストボックス128に「the change is」という言葉を挿入した。そして、エンドユーザは、適切な編集タスク機能バー86、88、90、92、94、および96をクリックすることにより、選択およびテキストを確認する。

【0053】図11は、ここではMarket Snapshotチャート114の選択部分のノードIDならびに挿入されたテキストを反映した、更新されたマスクエディタ80を示すことにより、エンドユーザマスク作成における次のステップを示す。エンドユーザカスタマイズプロセスのさらなる例は、図12に示され、ここでは、チャート114からボックス130がさらに選択され、新しいテキストがヘッダテキストボックス128に挿入される。マスクエディタ80は、ここで、選択されたボックスのノードIDを反映している。「select all」という編集タスク機能バー92をクリックすることで、チャート114の特定の列にあるすべてのテキストが選択され、これによって、エンドユーザカスタマイズプロセスが簡易化される。

【0054】図13に示すように、エンドユーザが作成したテキストは隣接するボックス130およびチャート114内に同様に配置されたすべてのボックスに現れる。マスクエディタ80は、この新しいテキストも反映している。エンドユーザがこのカスタマイズされたコンテンツ選択およびテキストを保存すると、次にエンドユーザは、選択されたウェブページにアクセスしたとき、「Market Snapshot」に含まれる情報が選択される（上述したキーパッドプロセスまたは音声コマンドプロセスを使用して）と、エンドユーザはアクセスしたウェブページ（図13では、特定の株式市場状況により例示される）から現在のテキスト、ならびに理解しやすいフォーマット提示に構成するために、カスタマイズされたテキスト（単に数字による株式市況からなる）の両方を受け取る。

【0055】説明した実施の形態は、カスタマイズに対してノードIDアプローチを可能にするために、タスク機能バーを利用したが、カスタマイズに特定のHTMLファイルを選択するために、「キーワード」を利用することも可能であることも意図されることを理解されたい。このノードID／「キーワード」アプローチは、エンドユーザがカスタマイズするHTMLページの所望の情報すべてをより精確に識別し、かつ見つけられるようになる。また、この発明の変換は、ウェブページの新しい視覚的な表示（オーディオまたは視覚的表示を有するオーディオを主としたものに対して）を作成するために

使用してもよいことを理解されたい。

【0056】このように、HTMLベースのボイスブラウザおよび変換システムの好ましい実施の形態を説明したが、当業者には、システムにおける特定の利点が達成されたことが自明であるはずである。また、その様々な変形、適応、および代替の各実施の形態を、この発明の範囲および精神内で行なうことを理解されたい。この発明はさらに、添付の特許請求の範囲によって定義される。

#### 【0057】

【発明の効果】以上のように、この発明によれば、エンドユーザが標準的な固定電話または携帯電話を介してユーザーがインターネットに一般的にアクセスできるシステムが提供される。

#### 【図面の簡単な説明】

【図1】 この発明の一実施の形態に従った、エンドユーザが電話機を介してインターネットにアクセスすることが可能なボイスブラウザシステムを例示する機能図である。

【図2】 この発明の一実施の形態に従って、ウェブページの自動変換がどのように動作するかの流れ図である。

【図3】 例示的なオリジナルのウェブページ、すなわちホームページウェブページの説明図である。

【図4】 ニュースセクションを含む特定のセクションを例示する、図3の例示的なウェブページのサブページの説明図である。

【図5】 ショッピングセクションを含む特定のセクションを例示する、図3の例示的なウェブページのサブページの説明図である。

【図6】 この発明の一実施の形態に従った、インテリジェントアクセスプロセスを例示する流れ図である。

【図7】 この発明の一実施の形態に従った、マスクエディタを備える図5の例示的なウェブページの説明図である。

【図8】 例示的なウェブページの説明図である。

【図9】 この発明の一実施の形態によりウェブページのカスタマイズが可能な、マスクエディタを備える図8の例示的なウェブページの説明図である。

【図10】 エンドユーザが挿入したテキストを反映する図9の例示的なウェブページの説明図である。

【図11】 図10のさらなるエンドユーザカスタマイズを反映する例示的なウェブページの説明図である。

【図12】 エンドユーザがさらに挿入したテキストを反映する図11の例示的なウェブページの説明図である。

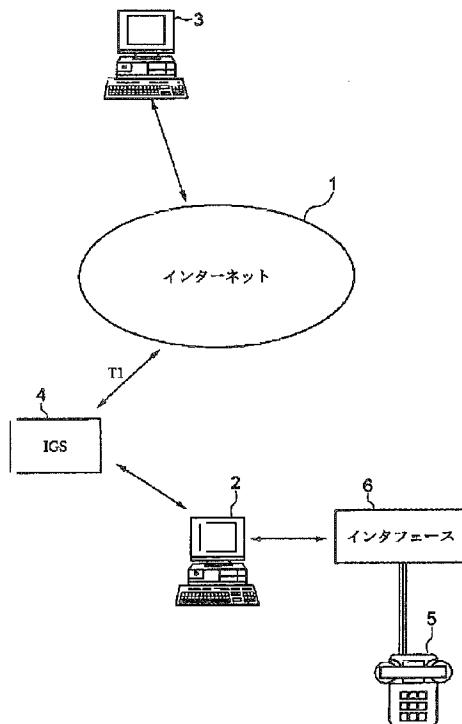
【図13】 図12のエンドユーザカスタマイズを反映する例示的なウェブページの説明図である。

#### 【符号の説明】

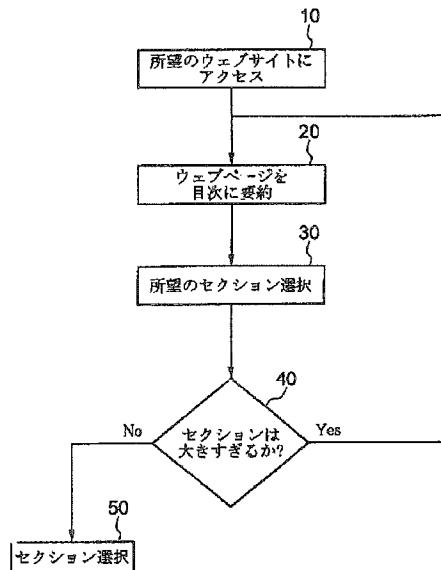
1 インターネット、2 ボイスブラウザクライアント

コンピュータ、3 ウェブサイトサーバ、4 インターネットゲートウェイサーバ、5 電話リンク、6 インターフェース。

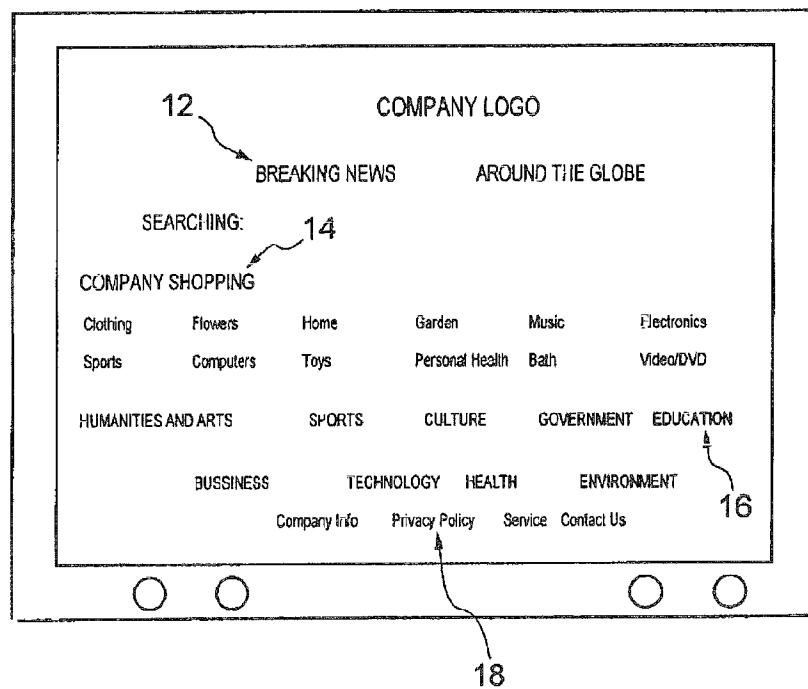
【図1】



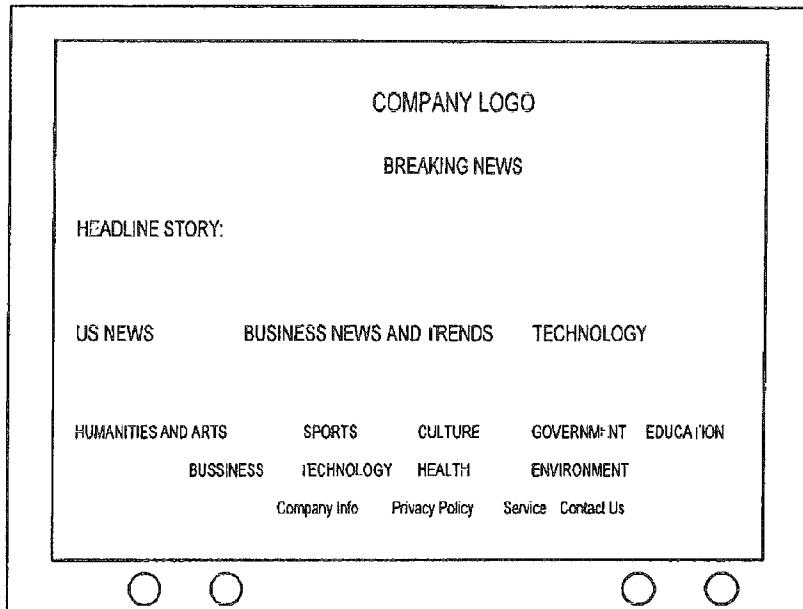
【図2】



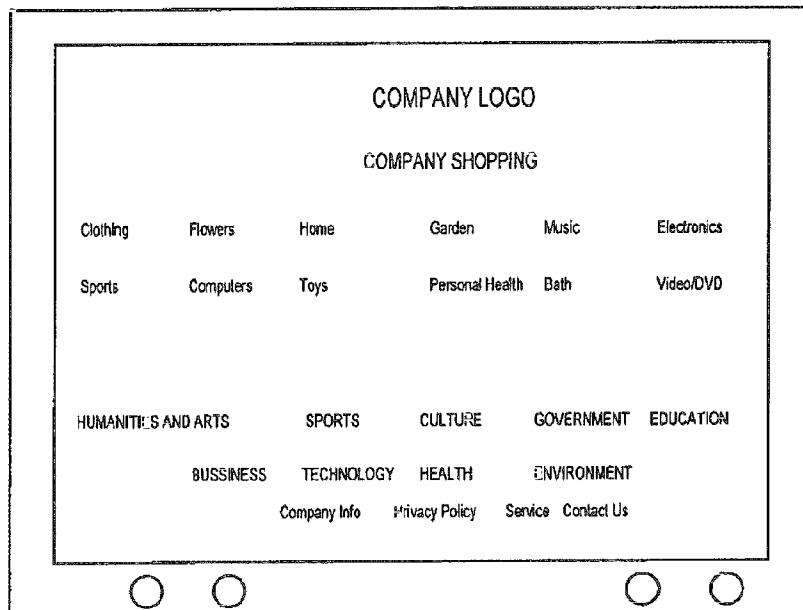
【図3】



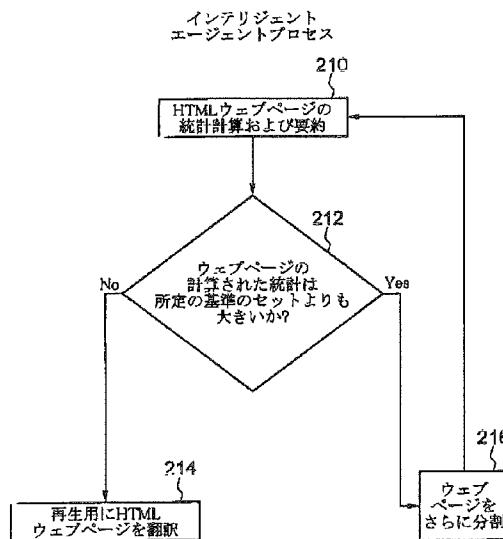
【図4】



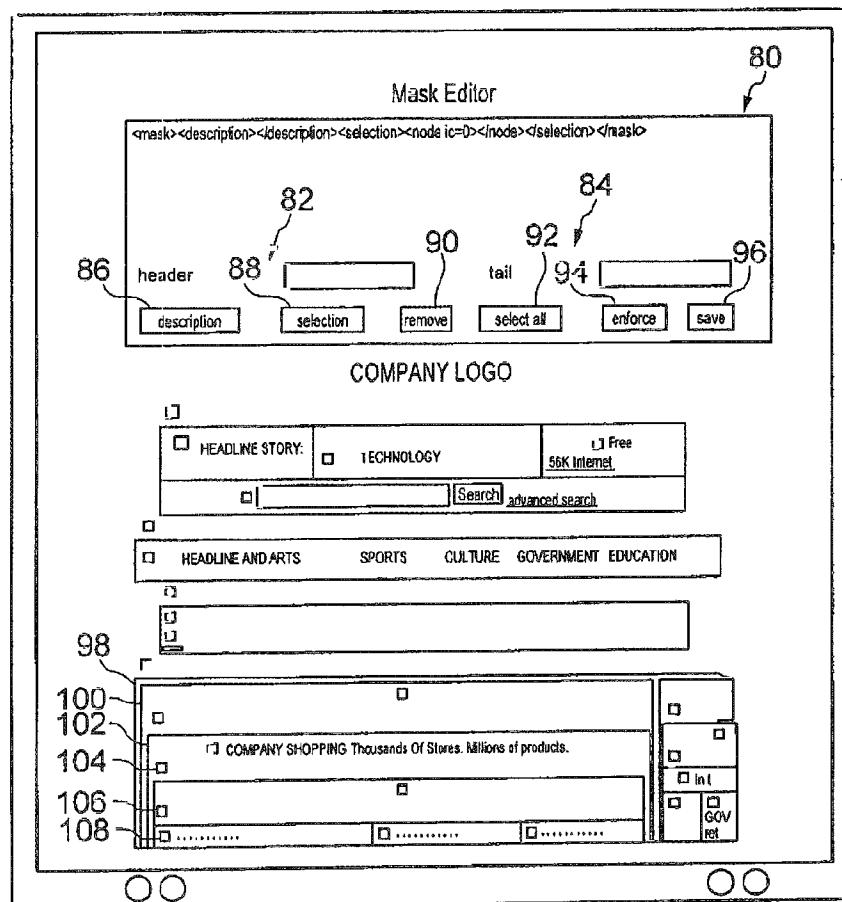
【図5】



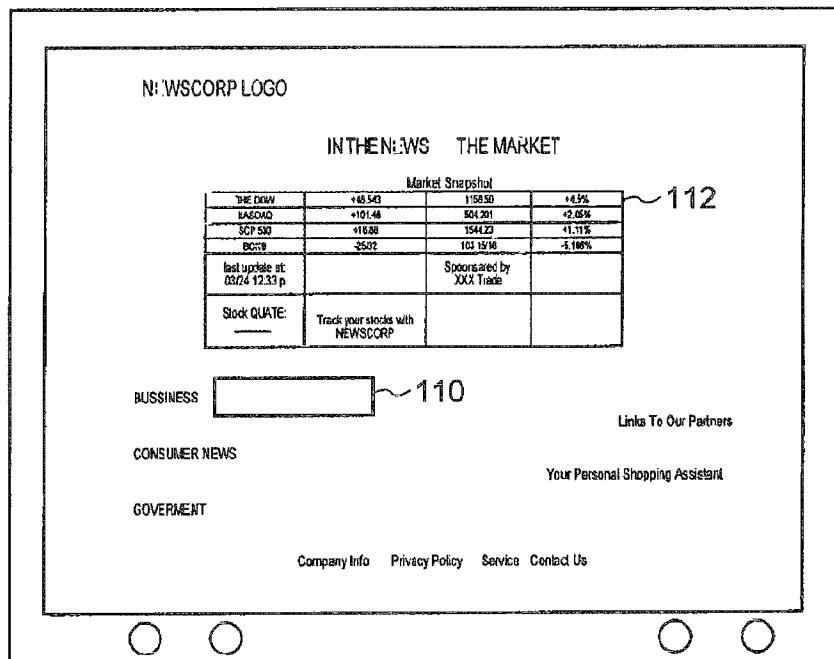
【図6】



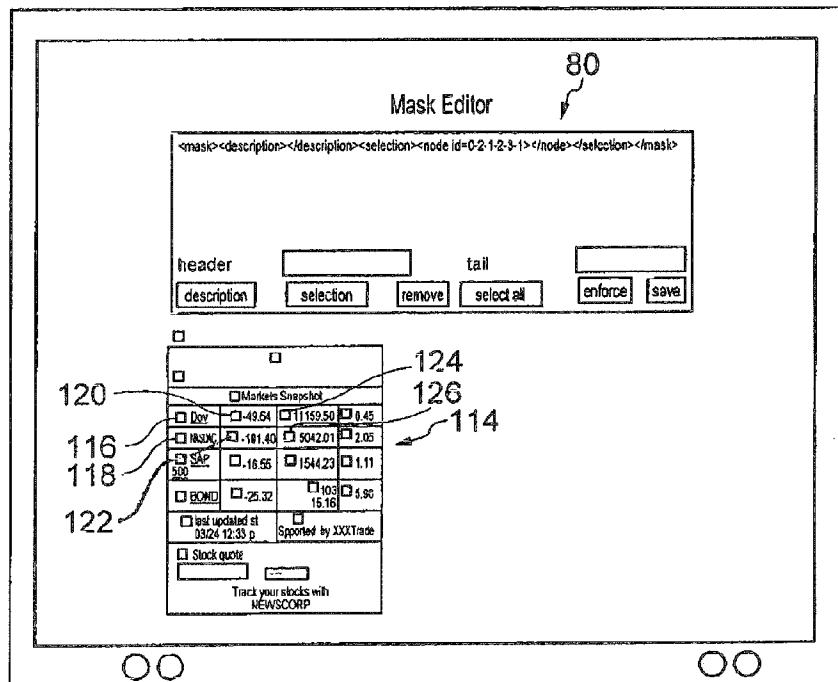
【図7】



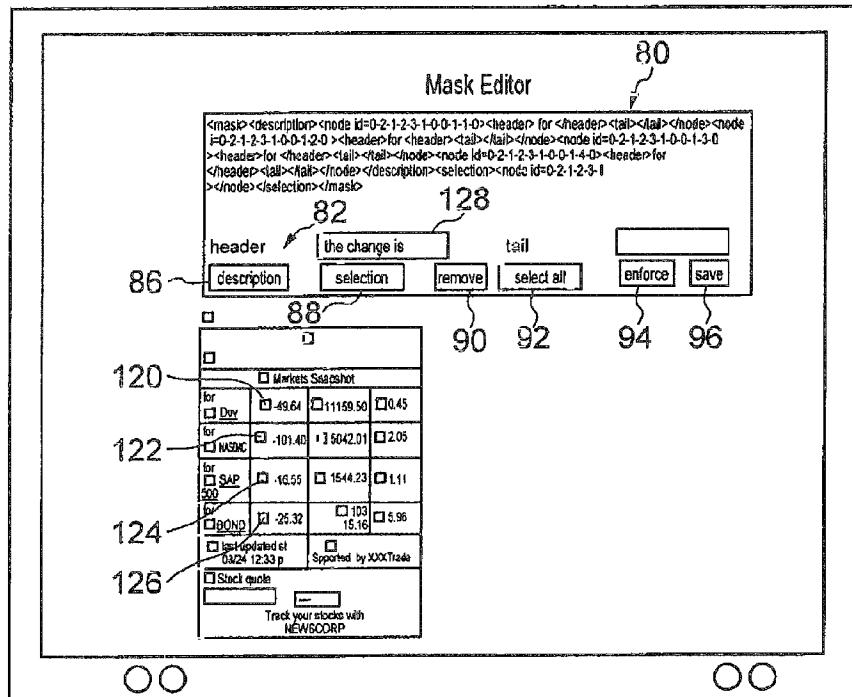
【図8】



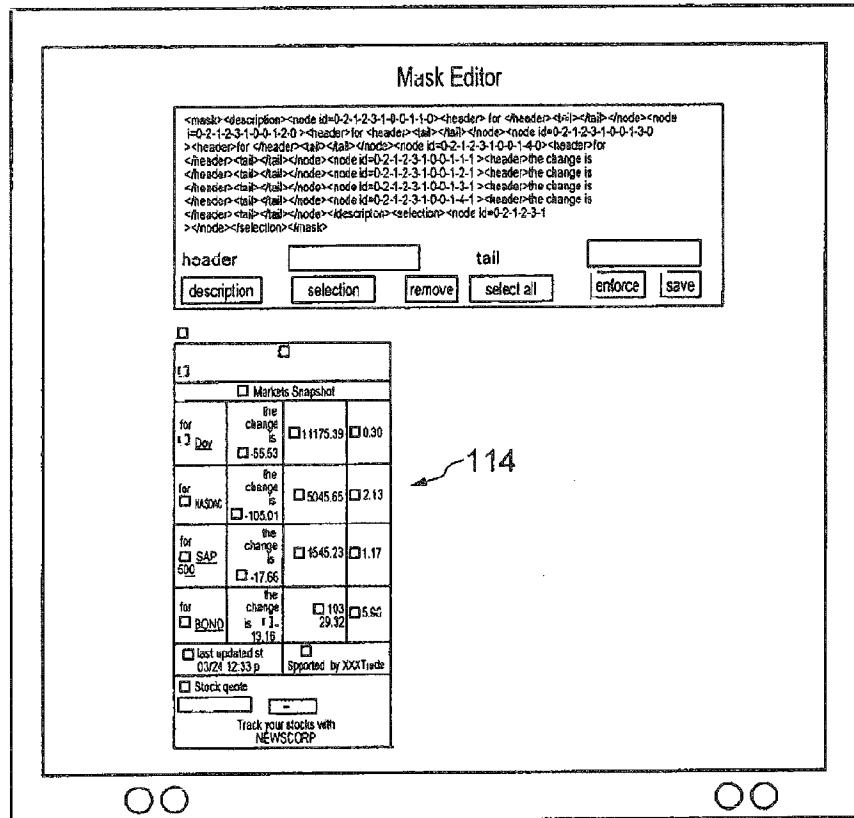
【図9】



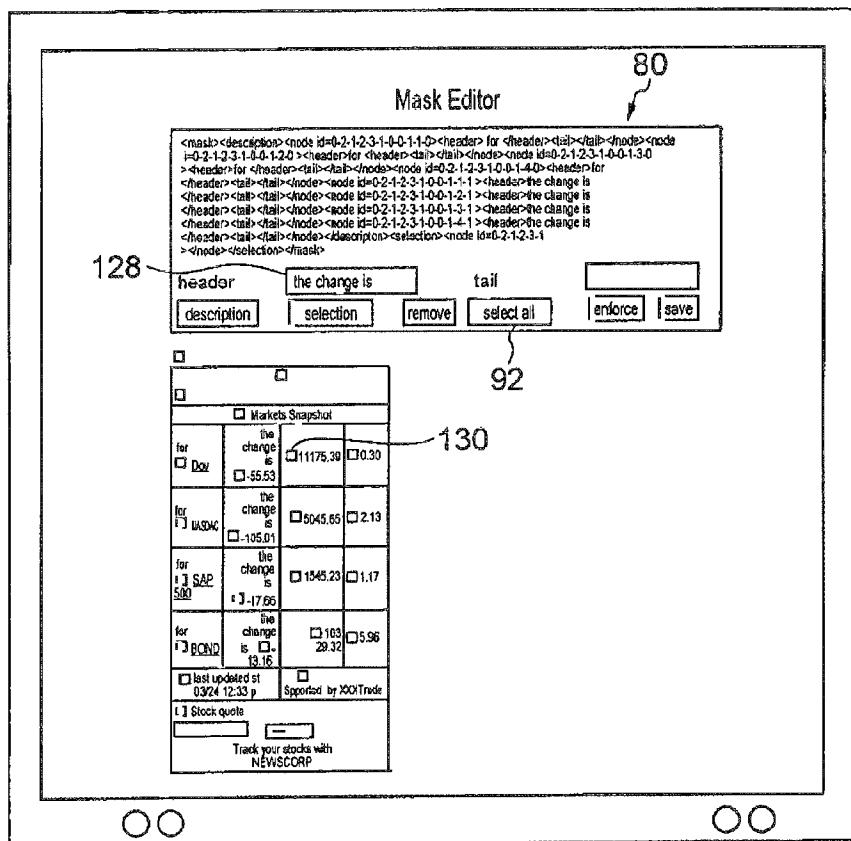
【図10】



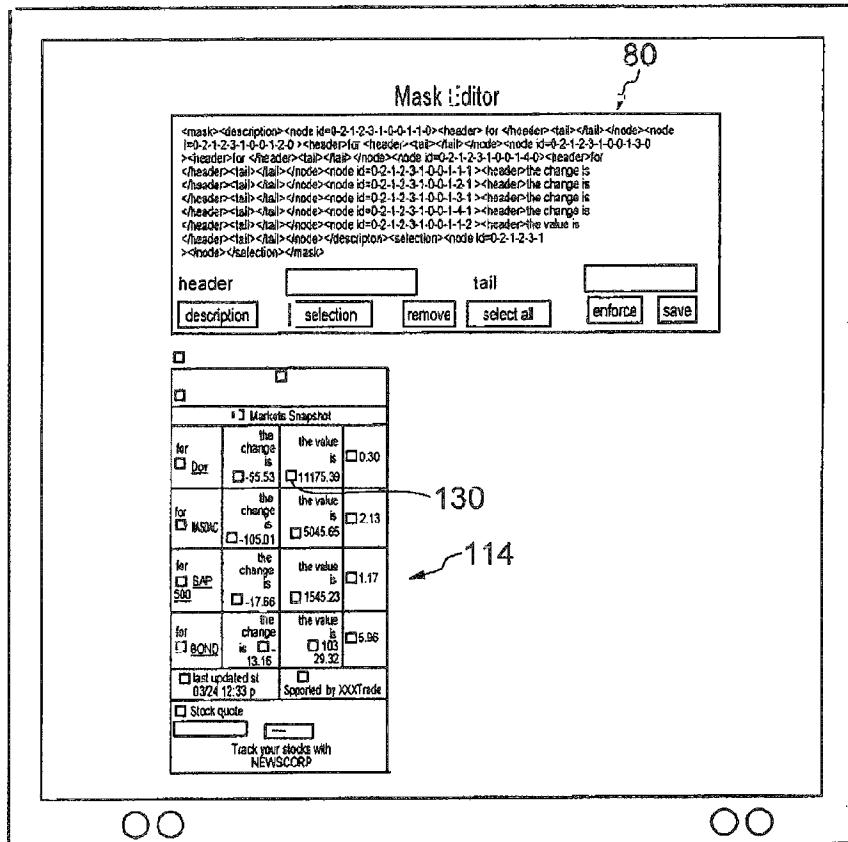
【図11】



【図12】



【図13】



## フロントページの続き

(71)出願人 500466566

128 Worcester Loop,  
Los Gatos, California  
95030, U. S. A.

(72)発明者 ジー・ワン

アメリカ合衆国、カリフォルニア州、ロサンゼルス、ソルトエア・アベニュー  
1409, #302

(72)発明者 ジイン・チャン

アメリカ合衆国、カリフォルニア州、パサデナ、サウス・アレン・アベニュー 159  
アパートメント 212

Fターム(参考) 5D015 KK02 KK03

5D045 AB01 AB26

## 【発明の名称】

ユーザがアクセス可能なウェブページを変換する方法、ユーザがアクセス可能なウェブページを自動的に変換するためのインテリジェントアーキャントプロセスを用いた方法 ボイスプラウ

(19) )01-331407 (P2001-331407A)

ザおよび変換システム、並びにウェブページにエンドユーザがカスタマイズしたマスクを作成する方法